Better to Organize Personal Information by Folders Or by Tags?: The Devil is in the Details

Andrea Civan
Medical Education and Biomedical Informatics, University of Washington, Seattle, Washington 98195
andreah@u.washington.edu

William Jones
The Information School, University of Washington, Seattle, Washington 98195
williamj@u.washington.edu

Predrag Klasnja
The Information School, University of Washington, Seattle, Washington 98195
klasnja@u.washington.edu

Harry Bruce
The Information School, University of Washington, Seattle, Washington 98195
harryb@u.washington.edu

People can organize information items by *placing* them into folders or by *tagging* them with labels. Over the years there has been considerable discussion on the relative merits of folders vs. labels. But there is relatively little empirical data directly comparing people's experiences with each model of organization on comparable sets of information.. We compared participants' experiences organizing information with both folders and labels over time. Results indicate that each model has its strengths and weaknesses when assessed with respect to basic activities of personal information management (PIM) such as keeping, organizing and re-finding. Results afford a deeper, more detailed understanding of each model in practice. This understanding is achieved by engaging participants in an active, "hands-on" comparison of models over a period of time. Study results point to syntheses in tool support for organizing information that might combine strengths of folders and labels while avoiding the inadvertent introduction of weaknesses from these models.

INTRODUCTION

Folders have long been the primary method for organizing personal information of various forms, ranging from paper and electronic documents to email messages and web references. But support for and interest in an alternate *tagging* model of information organization continues to increase. Are tags a passing fancy or are tags poised to supplant folders as the dominant means of organizing personal information? Alternatively, perhaps tags and folders will coexist. But, if so, how? In a "best of both worlds" combination, or in perpetual conflict?

These are important but, as yet, largely unstudied questions. Studies have explored the use of folders (or "directories") for web references (Abrams et al., 1998), email messages (Whittaker & Sidner, 1996; Bälter, 1997), and paper-based and electronic documents (Malone, 1983; Kwasnik, 1991; Barreau & Nardi, 1995). More recent studies compare the use of folders across document, email, and web bookmark organizations (Boardman & Sasse, 2004; Jones et al., 2005; Bergman et al., 2006). Other studies have explored the use of tags, particularly in web-based situations where people can benefit from the tagging efforts of others (Golder & Huberman, 2006; Kipp & Campbell, 2006; Ames & Naaman, 2007). However, Pak et al. (2007) note that while "*there is much enthusiasm about the purported benefits of tagging compared to more traditional ways of organizing information … this has not been supported by empirical evidence showing usability or performance benefits compared to traditional methods of information organization*"(p.5).

Pak et al.'s own pioneering comparisons of tagging vs. taxonomical systems for the organization of information yielded mixed results. In Experiment 1, which involved the organization of paper photographs, participants either "tagged" photographs by writing an unlimited number of keywords onto a Post-it note that was affixed to the photograph (tagging condition), or they organized photographs hierarchically through a combined use of Post-it notes and paper folders (taxonomic condition). Participants in the tagging condition organized photographs faster than participants in the taxonomic condition, even though they reported experiencing greater frustration. However, in Experiment 2, which

involved the organization of digital rather than paper photographs, participants were quicker to organize photographs in the taxonomic condition than in the tagging condition.

We face challenges in any attempt to compare models of information organization. For example, quantitative differences between comparison conditions, even when statistically reliable, may be inconclusive with respect to implications that might be reached -- either for the way people organize or for the design of supporting tools. In the Pak et al. experiment, for example, can we attribute discrepancies between Experiment 1 and Experiment 2 to some essential difference between organizing paper and digital photographs? Or could discrepancies be due to any of the many other details of system implementation that inevitably must be specified but that might limit the generalizability of results? For example, participants in Experiment 1 were able to sort or "pile" photographs in the taxonomic condition. While this may have been a satisfying activity (leading to less frustration), this activity took time. Was this the reason for the observed difference in organizing time between conditions? The devil, as the saying goes, is in the details.

A related challenge is to determine the essential nature of the comparison to be made. The tagging and taxonomic systems compared in Pak et al. map to points along a continuum: At one end are systems supporting highly structured (taxonomic, hierarchical) organizations; at the other end are systems supporting flat ("bushy") organizations. But the conditions in their experiment also differ with respect to a second dimension – photographs were organized by only one category in the taxonomic conditions but could have any number of categories ("labels") in the tagging conditions. Folders are commonly organized into a hierarchy whereas tags are often "flat" with no organizing constructs or interrelationships. However, a folder hierarchy can be broad rather than deep, and thus have a tag-like "bushiness" (i.e., many folders at the same level but no subfolders). We can also imagine a tagging system where labels are structured through constraints (e.g., pick "essential" or "optional" but not both), or organized into hierarchically structured facets (Hearst, 2006).

One essential difference underlies and inspires tool building efforts in support of folders vs. tags. Computer-supported folders follow a "put this there" model that agrees well with the way we organize objects in our physical world: Put the keys on the table; put the tax return in the "tax" folder. This model is simple and familiar, but also limiting. Information can be organized in only one way. An item of information can be in only one place at any time, even though we may have several ways that we'd like to organize our information (Kwasnik, 1991).

Tagging systems follow a different model. The information item isn't put anywhere and can be, in a digital sense, "placeless" (Dourish, 1999). Instead, the representation (labels) comes to the information item. An analog in the physical world might be a label gun used to affix any number of labels to a Tupperware container. But, as an important difference, digital tags can be searched and selected to see all associated items.

The use of tags would appear to have an obvious advantage over the use of folders: Whereas an information item is often in only one folder, it can have any number of tags. Tagging permits a many-to-many mapping (i.e., many tags to many documents), whereas folders permit only a one-to-many mapping (i.e., one folder can contain several documents). If a document is about both "apples" and "oranges" it can be tagged by both and accessed again later by either. But a document can be in and accessed from only one folder.

Given a choice then, why would anyone want to limit themselves to the use of a single folder when they could, instead, represent an item using any number of tags? One reason turns this limitation and apparent disadvantage of folders on its head. Yes, a document is in only one place but that means we know exactly where to look for it later. People are observed to prefer location-based methods for re-finding files on their personal computers (Barreau & Nardi, 1995), i.e., they select a likely folder and then browse within it to recognize a desired file. Jones (2007) reviews research in support of a more general preference for *wayfinding* methods that depend on a sense of digital location vs. direct search as the primary means for access to personal information. Indeed, Bergman et al. (2008) suggests that for many people direct search remains a last resort for retrieval of personal information only after attempts to "navigate" to information have failed.

However, we can think of exceptions and qualifications. Direct, keyword search through a search service (e.g., Google) is certainly a primary means of retrieving new, public information from the Internet. What about the retrieval of information items from a large database or collection (e.g. of articles) that may be personal in the sense that these reside on a personal hard drive but are otherwise new and not yet personally experienced? What about people, who by their nature or the unpredictable nature of their jobs (Bondarenko & Janssen, 2005), prefer to place most of their documents in unorganized "piles"? Whether people prefer direct search to browsing is likely to depend on a number of factors.

A similar truth could hold for the question posed in this paper's title. Better to organize information by folders or tags? The answer might very well be that "it depends". A first step, then, is to understand the nature of these dependencies more deeply – these "details" as it were. Armed with a better understanding of dependencies, we can answer additional questions: Under what circumstances and in what ways is it better to organize by folders or by tags? How would people like to be able to organize their information – if only they could?

This paper describes an exploratory study to gain qualitative insights into the relative advantages and disadvantages of two basic models of information organization and retrieval: Placing into folders vs. tagging with labels. In particular, how do these models compare in situations of personal information management (Jones & Teevan, 2007), where people are working on their computers and in their workspaces? How do these models compare over time as older information becomes "more personal" through familiarity? … as new information continues to arrive and the corpus of information to be managed increases in size? … as a person's understanding of this information and its uses evolves?

THE STUDY

The goal of this study was to understand how people go about organizing information, over time, using folders and labels. This study asked: In what ways are placing into folders and tagging with labels alike and different as models for personal information management? Several aspects of the study are worthy of note:

- *The study was exploratory*. Although quantitative measures of performance were taken, the emphasis was primarily on qualitative and in-depth comparisons expressed by participants. As Malone (1983) notes, the value of such a study is often in the insights and compelling examples that it provides. In some cases, examples and insights deserve special focus in follow-on studies. In other cases, examples and insights may point directly to implications for tool design.
- *The comparison was "within subjects"*. The two conditions, tagging with labels vs. placing into folders, were experienced by all participants. Moreover, participants were actively involved and were encouraged to reflect upon and record (in a daily diary) their perception of the differences between these two conditions.
- *The manipulation was achieved by using two actual systems* of email information management: The basic web-based versions of Microsoft's Hotmail email service (http://get.live.com/mail/classic_features2) and Google's Gmail email service (http://www.gmail.google.com/). As discussed below, the user interfaces for these two systems are similar in many respects. However, the systems differ with respect to their underlying model for information organization and retrieval. An email message in Hotmail is placed in and can later be accessed via a folder. A message can be in only one folder. By contrast, an email message in Gmail is tagged with any number of labels and can later be accessed through any of those labels. Participants were encouraged to observe and comment on this essential difference between the systems.
- *Participants worked over time with the information for both conditions on their own computers and in their own workspaces*. Information in each condition related to practical projects of the participant's choosing. The information for each condition arrived in email messages over a period of five days.

Selection of Systems for Comparison

Clarity concerning the nature of the targeted comparison – placing into folders vs. tagging with labels – does not directly point to a means for achieving this comparison. We considered prototyping two systems of organization that were similar in all respects except for the model of information organization – placing vs. tagging. However, it quickly became clear that even basic systems could not be implemented without considerable expense. Even after the investment of time and programming resources, the resulting systems might still be "buggy" and lack essential features, which might distract from the targeted comparison. We elected, instead, to engage participants in a comparison using two existing web-based email services, Hotmail and Gmail. Each service has been tested and refined through years of experience with tens of millions of users. The versions we used were equally rudimentary in their user interface. Most differences between these systems are minor and could be overlooked by participants for the purposes of this study. For example, Hotmail has a "Junk" folder and Gmail has a "Spam" label. However, Hotmail and Gmail differ considerably in one essential respect: Users of Hotmail place email messages into folders whereas users of Gmail tag email messages with labels (Figure 1).

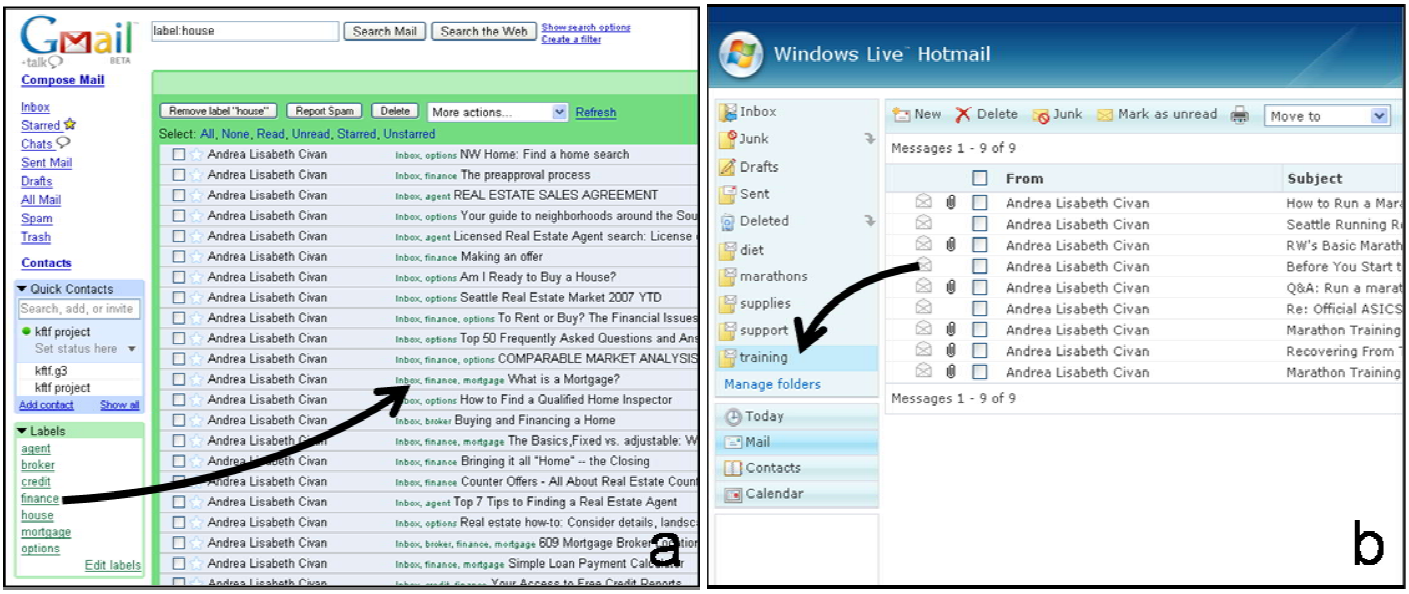

Figure 1. Gmail (a) and Hotmail (b) differ in their models of information organization. Hotmail supports a traditional model of placing messages into folders. In Gmail, labels are used to tag messages that stay in place.

Given our choice of systems for the comparison, we emphasize that this study is not about email systems or a comparison of usability. Instead, Hotmail and Gmail provided a convenient basis for comparing placing and tagging models of information organization. The email transport in both systems provided an important control of information delivery. In our everyday lives, the information we encounter rarely arrives in one neatly organized package. Instead, it arrives piecemeal, over time. As new information arrives, our ways of understanding a topic and creating a scheme to organize associated information will often change. Similarly, we used Hotmail and Gmail to send participants small collections of materials in piecemeal packets over time. After using both models to organize this information, each participant reported on how their experiences compared. Certainly, some differences identified by participants would be obvious. We expected, for example, that participants would note the flexibility afforded by the tagging model. Other differences could reflect a comparison between Hotmail and Gmail services that have little directly to do with the comparison of placing vs. tagging models. There is then a third category of observations – observations relating to important but non-obvious differences between the two models of information organization. These differences were evident to participants, but only after a period of hands-on use of the two systems.

Methods

The comparison condition (i.e., placing into folders vs. tagging with labels) was used as a springboard to engage participants in targeted, situated discussion about the similarities and differences they experienced between organizing information with folders and labels. Each participant completed three study components: (1) a 30-minute initial interview, (2) a five-day information organization task using folders and labels, and (3) a one-hour follow-up session. The study produced interview transcripts, daily screen shots, participant diaries and sketches, and measures of speed and accuracy on retrieval tasks.

Procedures

During the *initial interview*, the participant described their previous experience using folders and labels to organize personal information, including tools they have used. They discussed what they like and dislike about folders and labels, what they find easy and not easy about each, whether they have a preference for one or the other, and how they view these models of organizing information as similar and different.

During the *information organization task*, the participant gained hands-on experience with folders and labels as they organized project-related materials using Hotmail and Gmail over five days. Materials were web articles related to ten

project topics. Each topic was a popular personal goal (see Table 1). For each topic, we constructed a collection of 25 articles. We used the recommended steps for completing each project from a popular "how to" website (http://www.soyouwanna.com/) to guide our selection of relevant articles. The participant selected two from the ten possible projects. They were instructed to select topics they were less familiar with, but that would keep their interest over five days as we emailed them batches of articles to organize.

Both a Hotmail and a Gmail account were created for the participant to use in this study. On each of the five days, we sent a batch of five articles from the first project collection to the Hotmail account (i.e., "placing" condition) and a batch of five articles from the second project collection to the Gmail account (i.e.," tagging" condition). Effort was taken to balance the assignment of projects between conditions. For example, when two participants selected the same project topic, we assigned that project to Hotmail for the first participant and to Gmail for the second.

During each of the five days, the participant spent 10-15 minutes organizing materials they received in Hotmail using folders and in Gmail using labels. We asked them to keep a brief diary about their experiences using folders and labels. We also took daily screenshots of the Hotmail and Gmail account to observe how their organizational schemes evolved.

Within seven days of completing the information organization task, we conducted a *follow-up session* to learn about the participant's experiences using folders and labels to organize information. First, we measured the participant's performance on retrieval tasks. For each condition, we asked the participant to:

- Recall from memory three specific details about the project based on the articles they organized
  We randomly selected one from the five recommended steps for completing the project suggested by soyouwanna.com. Without looking back at their project collection, the participant wrote down up to three details about that step of the project. We measured the number of details recalled out of three (i.e., 1 point for each detail that was present in an article, ½ point for each relevant detail that was not present in an article, and no point if they wrote down no detail).

- Re-find five randomly selected articles based on verbal cues
  We asked the participant to re-find each article after providing them with a brief verbal cue (i.e., *"the article about national parks in Australia*"). We measured the time it took the participant to re-find the cued article, the number of places they looked (e.g., individual folders/labels, the inbox), and whether they located the cued article.

Next, the participant drew a free-hand sketch to show how they conceptualize the information they received for each project as a whole. The sketches were used to walk us through the organizational schemes they created with folders and with labels. Finally, we asked the participant to reflect on their experience using folders and labels to organize information by repeating the questions we asked them during the initial interview.

Following data collection, we created a case report for each participant, which included qualitative views, quantitative performance measures, and visual mappings that represented the associations they made between articles they received and the folders/labels they created to organize those articles over time. We transcribed and coded interviews and participant diaries for emergent themes. We calculated mean scores for performance on retrieval tasks. We created visual mappings from screenshots to track the evolution of their organizational schemes over time, which we describe further in Results. Finally, we synthesized qualitative themes, performance scores, and patterns in visual mappings across cases.

Participants

Participants were four women and six men who ranged in age from 19 to 35 years (average age 25). All were students in health-related and informatics-related fields at undergraduate (N=5) and graduate (N=5) levels. Each had routine experience with folders and had used labels before, typically with online tools (e.g., Gmail, del.icio.us). Six participants used these systems daily (i.e., "routine tag-users"). Table 1 shows the number of participants who selected each project by condition.

Project topic	Placing	Tagging
1. Publishing a book	1	2
2. Writing a business plan	1	0
3. Getting a pilot's license	0	1

4. Taking up yoga	1	0
5. Learning to scuba dive	1	0
6. Preparing for a marathon	2	1
7. Lowering your cholesterol	1	1
8. Buying a home	1	2
9. Planning a trip to Australia	1	1
10. Setting up a savings budget	1	2

Table 1. Project selection and assignment to condition

Three participants expressed a general preference for folders and two did so for labels. The remaining five indicated that their preference depended on situational factors, such as format or task. For example, one participant told us: *"For emailing and my desktop, for archiving email, I like folders better. Tagging doesn't necessarily work for that, folders work better. But in different situations tags can be very useful, like in grouping music into different genres*."

RESULTS

We first describe similarities in the ways participants organized information with folders and labels, including common problems they experienced with each model. We then describe important differences between folders and labels, including pros and cons of using each.

Similarities between Placing and Tagging

The two models of information organization were comparable with respect to:

- Retrieval performance
- Evolution in mappings between articles and folders/labels over time
- Limitations to fully express one's internal conceptualization

Parallels in Retrieval Performance

No differences were apparent in performance on retrieval tasks between placing articles into folders and tagging articles with labels (Table 2). Participants recalled roughly the same number of details from memory in both conditions. They were able to re-find every cued article in both conditions quickly by looking in one or two places.

Table 2. Mean performance on retrieval tasks

Performance Measure	Placing	Tagging	Difference (paired t-test)
Details recalled (out of 3)	1.9	2.4	p = 0.30
Time to re-find cued articles	14.5 s	15.4 s	p = 0.84
Places looked to re-find cued articles	1.5	1.4	p = 0.59

Both Hotmail and Gmail provide keyword search, which participants could have used instead of browsing (i.e., clicking on a label/folder, then scanning the resulting list). One participant first used search to re-find all cued articles in both conditions, reverting to browse only for two labeled articles after the search had failed to return them. The other nine participants used browsing exclusively as their retrieval method.

Parallels in the Evolution of Organizational Schemes

The arrival of new articles over successive days prompted participants to create new folders and labels. New articles sometimes prompted a revision of the organizational scheme that mapped folders/labels to articles in which older folders/labels were deleted and their associated articles re-assigned. Regardless of condition, these mappings grew and changed over time to accommodate the arrival of new information, as well as changes in the participant's understanding of this information.

In order to compare changes in the mappings between articles and folders/labels in each condition, the screenshots collected on successive days from each participant were used to create visual mappings, such as those shown for Participant "P2" in Figure 2. Visual mapping "a" (Figure 2, left) shows how P2 mapped articles about "Preparing for a marathon" to the folders he created. Visual mapping "b" (Figure 2, right) shows how P2 mapped articles about "Publishing a book" to the organizational scheme he created with labels.

In each visual mapping, the circles in the left column represent the twenty-five articles that P2 received over 5 days. The rectangles in the right column represent the folders (Figure 2a) or labels (Figure 2b) that P2 created. Assignment of an article to a folder/label is shown by a solid line between a circle and a rectangle. Articles, folders/labels, and lines are color coded according to the day on which articles arrived and were assigned. A dashed line represents an article that P2 had assigned to a folder/label, but subsequently removed. We placed an "x" mark over a folder/label that was created, but subsequently deleted. Dashed lines and "x" marks indicate reorganization of the organizational scheme over time.

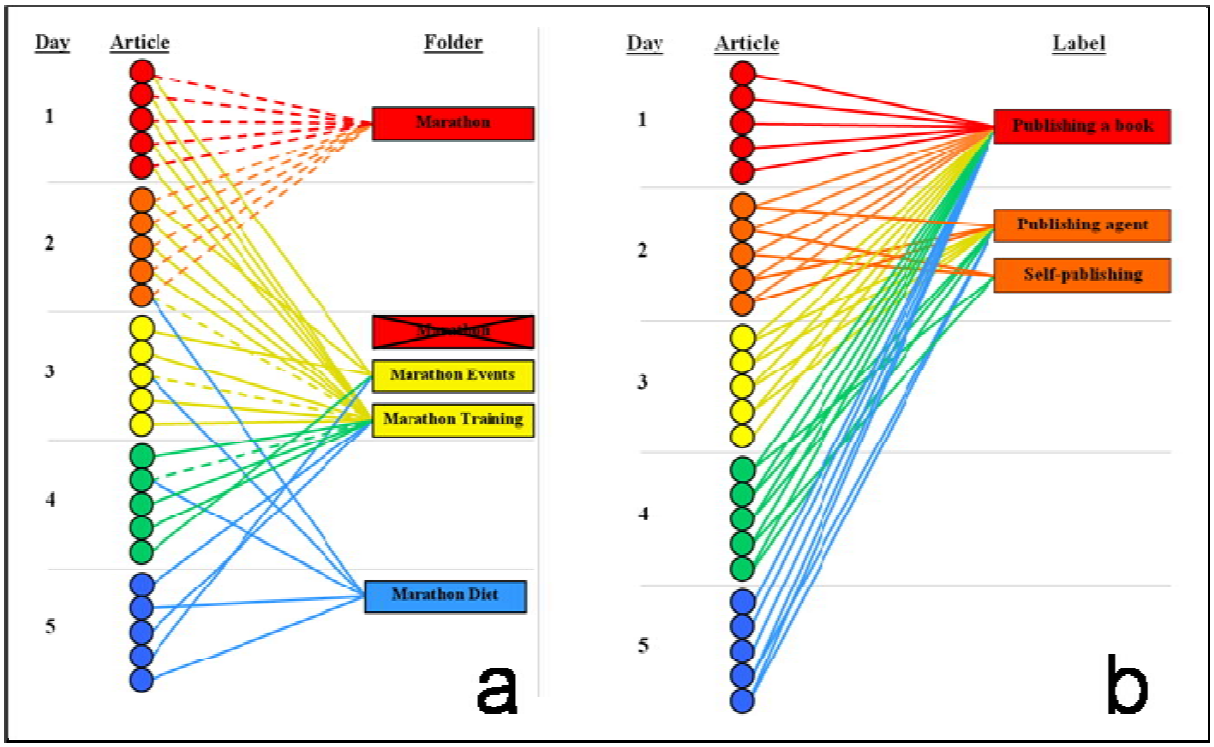

Figure 2. Visual mappings show changes over time in the mapping between articles and folders (a) and between articles and labels (b)

Taken together, the participants' visual mappings suggest that the organizational schemes they created with folders and labels followed similar patterns of growth and development. The rate with which folders and labels were created was similar across participants. On average, participants applied one more label than folder to the batch of articles they received each day. On the final day, they associated a similar number of articles with each folder (6.2 articles/folder) and with each label (7.2 articles/label). On average, they assigned 1.4 labels per article.

Although neither Hotmail nor Gmail explicitly supports hierarchy, six participants used work-arounds to incorporate whole-part and general-specific relationships into their organizational scheme using Hotmail and seven did so using Gmail. Common work-arounds in both conditions included naming conventions, such as grouping non-specific articles into a folder/label named "General" or adding a common prefix to multiple folders/labels (e.g., Figure 2a shows three folder names that begin with "Marathon"). Although the capacity to establish multiple categorizations distinguishes labels from folders, only five participants created many-to-many mappings in the organizational schemes they created with labels. For example, Figure 2b shows how P2 associated individual articles with multiple labels. Four of the five participants who used multiple categorizations were routine tag-users. Some participants who decided against multiple categorizations voiced concern that utilizing them could cause confusion, redundancy, and inefficiency: *"At first I liked*

that I could put each one [article] *in more than one tag, but later as I organized the emails, I thought that maybe that would just be redundant information for me when I am looking for the information later. So I didn't end up putting more than one tag on emails."*

Parallel Limits on External Expression

The organizational schemes participants created for their project collections, whether using folders or labels, were considerably less rich than what they could express through their hand-drawn sketches. These sketches reflect the participants' understanding of their collections (i.e., external representations for how they conceive of their information and how items relate). For example, Figure 3 shows the sketches that P2 drew. Sketch "a" (Figure 3, left) represents the project "Preparing for a marathon", which P2 organized with folders and corresponds to the visual mapping shown in Figure 2a. Sketch "b" (Figure 3, right) represents the project "Publishing a book" and corresponds to P2's visual mapping for labels (Figure 2b). Although the visual mapping and sketch for the label condition are quite similar (compare Figure 2b and Figure 3b), P2's sketch for the folder condition (Figure 3a) is much more elaborate than what he could realize using folders (Figure 2a).

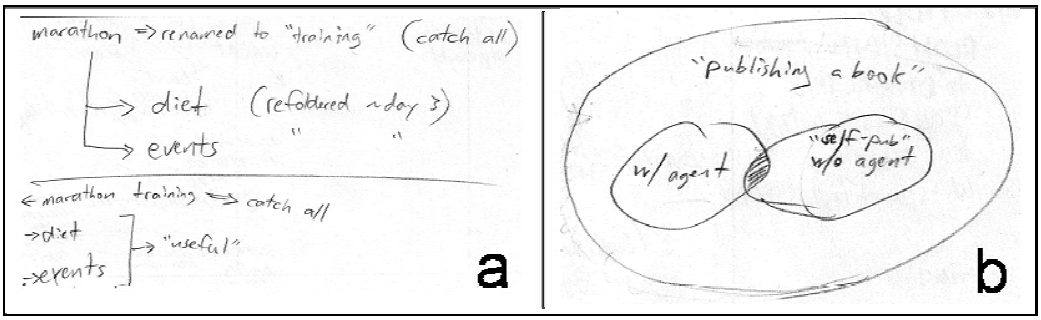

Figure 3. Tree-like sketch for a collection about "Preparing for a marathon" organized with folders (a) and Venn diagram-like sketch for a collection about "Publishing a book" organized with labels (b)

However, a close alignment between sketches and visual mapping for labels was not the case for all participants. For example, Figure 4 shows the much more elaborate sketch Participant "P4" drew for the same collection, "Publishing a book", which she also organized with labels (compare to P2's sketch in Figure 3b). The arrow shown at the top of Figure 4 indicates P4's conceptualization of the temporal flow of this project's tasks, which is a detail she was unable to represent using labels.

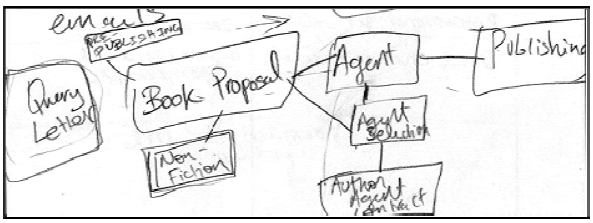

Figure 4. Workflow-like sketch for a collection about "Publishing a book" organized with labels

Comparisons between participants' sketches and the organizational schemes they created, as depicted by their corresponding visual mappings, highlight the limits of both folders and labels for expressing the rich details from internal views that participants held about their project collections. For example, six participants incorporated hierarchical relationships into the sketches of their folder collections and seven did so for their tag collections. Two participants incorporated temporal characteristics into the sketches of their tag collections. However, neither use of folders nor labels allowed participants to explicitly incorporate these kinds of details into their organizational schemes.

Key Differences between Placing and Tagging

Participants expressed an important difference between placing items into folders and tagging items with labels with respect to perceived control. Most felt that folders provided control over their information by helping them to get items out of their way: *"The emails in the inbox were gone. I felt more in control because of that. I felt like I was dealing with my emails."* Although hiding information was a key advantage that participants associated with folders, two participants used Gmail's archive function to gain control by getting labeled items out of their inbox. In contrast to folders, most participants felt that labels provided control by offering freedom to keep and re-find information in multiple ways: *"There was a lot more control with tags because you can apply multiple tags. You can refer to something with lots of different names and with folders it just sits in one place. The control comes from the flexibility of being able to reference it in multiple ways."*

More specifically, key differences between folders and labels surfaced for keeping, organizing, and re-finding information. Although both folders and labels helped participants to manage their information, these differences suggest unique advantages that each model offered. Next, we delineate these advantages, as well as drawbacks, by contrasting participants' views and uses of folders and labels both on the "input side" -to keep and organize information within an organizational scheme, and on the "output side" -to later re-find that information for re-use.

Differences on the Input Side: Keeping & Organizing Information

Cognitive versus Physical Effort

We identified two costs of keeping and organizing information, each associated with a particular form of effort put forth by participants. First, it required cognitive effort for participants to select an appropriate category for an article. Second, it required physical (i.e., manual) effort for participants to assign that article to either a folder or label. Although most participants reported spending a similar amount of time organizing articles with folders and labels, the balance between physical and cognitive effort played out differently between these two models.

One participant described the high cognitive effort required to select each and every relevant label to tag an article. In contrast, the other nine participants agreed that it required more cognitive effort (e.g., "*choosiness"*, "*commitment"*) to decide which folder, rather than which label(s), best fit an item: "*With tags it was easier to work through and pick more tags. But with folders I kind of had to think, 'hmm does this really fit in this category?', instead of saying 'of course it fits in this one and in that one too'. I had to choose to make sure that I wouldn't lose it, but with tags*, *I had the option to be less choosy."*

Participants agreed that it took slightly more physical effort to tag an article with multiple labels than to place an article into a single folder: *"I felt like I spent more time on the tags just because I had to apply multiple tags and so I had to apply one and the next one and the next one. With folders, I just had to do it once."*

Hiding Information versus Keeping Information Visible

Although visible information captures our attention and helps us to stay on task, clutter from too much visible information disrupts our attention and tasks. Hiding information, by moving items out of our way, reduces clutter but can also cause us to forget where we put items. Folders and labels offered participants different benefits in the tradeoff between hiding articles they did not want to look at immediately and keeping visible articles they wanted to get back to later.

Most participants found it was easier to hide information with folders than with labels: *"Folders hide information. I am trying to hide some information so that I can more easily access the information I want. I don't need to look at that information all the time. It actually requires cognitive effort for me to go through and look at it, and read this title and say 'Oh, I don't need to look at it'. But if I need it, I know that it is being hidden and I know where I can go look for it*." Most participants also agreed that hiding articles in folders could also cause them to forget: *"With folders, I don't like that things get hidden away and then I might miss something important."*

Two participants preferred not to hide any articles they organized with labels: *"I really wanted to browse all of my messages in a succession easily, so have everything in one box and then be able to see things from there. And if I used*

folders, then everything would be lumped into sections." However, the remaining eight participants expressed frustration because tagged articles did not leave their inbox: *"Sometimes I want to hide information* [in Gmail]. *Visually, I can only sort through so much stuff at one time."* Two participants hid articles using Gmail's archive function: *"I really like Gmail for the archive folder. It really helps keep it* [inbox] *clean so I know what I am dealing with."*

A different pattern emerged in the ways participants kept information visible using folders and labels. Most participants highlighted articles using folders by keeping them uncategorized in their inbox, which often provided a "to do" list: *"If you put something in a folder and you need to do something with it, then it is not sitting there reminding you, it may be hidden. I have a tendency not to put it in a folder if I need to do something with it".*

Several participants also kept an article temporarily in their inbox while they decided which folder it best fit. This visibility of uncategorized articles offered a surprising advantage over labels. Eight participants kept all tagged and untagged articles in their Gmail inbox. They described frustration because they were unable to easily locate only untagged items: "*if a message isn't tagged, I don't know how to find it. With tags, it's hard to get at only the uncategorized stuff."* Two participants highlighted articles with Gmail's star label, making their ongoing tasks more visible: *"The star means that I need to do something with it -a reminder for me to go back, a quick and easy way to highlight something that I need to get back to."*

Differences on the Output Side: Re-finding information

Flexible vs. systematic searching

Several participants found folders better for searching all possible routes back to an article when they forgot the exact folder name where they had placed it. Labels were viewed by most participants as more limited for this systematic, exhaustive search process. For example, one participant expressed frustration that she could not combine labels: *"I just want the 'AND' combination."* Another participant noted: *"You can exhaustively search a set of folders, you can search one folder and if it's not there, you search the next folder and so on. But you can't do the same thing with labels because they overlap so much and you don't want to look at the same things over and over again. That makes search inefficient and redundant."*

However, most participants found that labels provided greater flexibility than folders by offering multiple paths back to an item: *"Tags are useful if email is about two topics. I can assign it with two unrelated categories and it will show up under either search. But with folders you can only assign things to one folder, so you have to know the exact folder name to find it."*

Re-finding Cues Offered by Folders and Labels

Participants described different kinds of cues that folders and labels offer for re-finding. For example, folders can provide visual cues that allow for the use of spatial memory and recognition in re-finding: *"Folders are visual, I like that you can just sort of glance at it and remember it by where it is located."* However, participants appreciated the value of labels not only as a means of access to information but also as a quick summary of article content and category: *"I liked that I could see the tags in front of the email because I can right away see what kind of information they are."* Another participant noted: *"Having Gmail show me a little more info about everything is a lot easier than reading every title."*

The informational cues of labels coupled with their multiple retrieval paths offer opportunities for serendipitous encounters, whereas this advantage was not associated with folders: *"I like tags when you are trying to find something that's similar to something I already have, but different than that exact thing."* A few participants noted that labels made it easier than folders to discover content by revealing unexpected or forgotten connections between articles: *"It is a lot easier* [with labels] *to jump to ideas, or content that I recalled or want to remember instead of just clicking and going through them all* [in folders]*."*

CONCLUSIONS

Folders or labels? Placing or tagging? Our results suggest the need for a little of each and more than either. This study helped to identify and clarify differences and similarities between folders and labels and, more generally, between placing and tagging as models for information organization.

The noted differences do more to identify tradeoffs and tensions than to identify a clear winner between models. On the "input side", for example, folders provided an effective way to manage workflow and inbox clutter. Once an article was placed into a folder, it was no longer in the inbox. The remaining articles were those that still needed to be processed. Although participants acknowledged the ability to "place" articles using folders, they also noted that articles could be "misplaced" into folders, which could later lead to longer retrieval efforts or forgotten articles (i.e., "out of sight, out of mind").

The actions of tagging were substantially different. Even after an article was labeled, it stayed in the inbox. Although some participants liked having all articles in "*one place*", enabling a complete scan, they were forced to take explicit steps to achieve control over workflow and clutter comparable to that of folders. For example, participants starred articles that still needed work or they gave articles a workflow label like "to read". Once articles had been processed, these labels then needed to be removed. One desirable feature suggested by participants was filtering not only by a given label, but also by the absence of labels (i.e., show me only articles that have no label).

Although participants appreciated the use of folders for one kind of control (i.e., moving articles out of the inbox into folders), they also liked labels for a different kind of control: an article could be labeled in several different ways rather than being placed in only one folder. Here too, however, participants noted a tradeoff. Picking the "right" folder took more thought (i.e., "cognitive effort"). On the other hand, to give an article several labels took more manual effort since the same action of tagging had to be repeated.

Tradeoffs were also noted on the "output side" during retrieval as articles were accessed later. For example, participants appreciated the option of accessing articles through several different labels, which could uncover unexpected and otherwise hidden connections between articles. However, when participants were not sure which of several folders or labels might lead to a desired article, a "try here then try there" search worked better for folders than for labels. When doing such a systematic search in the label condition, participants experienced frustration because articles eliminated from consideration during the scan of one label's listing of articles would recur in listings for other labels.

Similarities between the two conditions are also noteworthy. First, consistent with previous research (Barreau & Nardi, 1995; Bergman et al., 2008) our participants preferred to re-find articles in both conditions by browsing. Second, although neither Hotmail nor Gmail provided support to order folders or labels, or to organize articles according to part/whole or general-specific relationships, participants still found ad hoc ways to do so (e.g., through their selection of folder and label names). Lastly, participants were able to quickly sketch by hand much richer representations in which concepts were related, for example, by time and workflow dependencies.

We emphasize two practical implications of the study's results:

- Don't leave the good stuff behind! In a shift to a tagging model of information organization, care should be taken to preserve the benefits of folders. Part/whole or general/specific relationships can easily be expressed using a folder hierarchy, for example. People can "trick" a tagging system to realize these hierarchical relationships, through naming conventions, for example. A tagging system that allows subsets of tags to be related within a hierarchy could make these relationships easier to express. Similarly, subsets of tags might be related to one another in a mutually exclusive/collective exhaustive relationship to achieve a partition of information that happens, de facto, when information is organized into folders.
- Better support for information organization may need to go well beyond folders and tags or their artful combination. As revealed in participants' sketches, people think of their information in ways that go well beyond the representational ability of either folders or tags. Participants appeared to organize information internally, for example, with respect to time and the steps of a workflow. These internal organizations evolve over time. How can our tools better support us so that corresponding external representations can stay in synch? What if, for example, people could use a digital "sketch pad" to create, expand, and refine the organization of an information collection over time? The ultimate model of information organization may be neither "place this" nor "label this", but instead, "this is how I see things".

Which is better? Folders or tags? Each model of information organization has its advantages and disadvantages – the devil is in the details. Participants, who are actively engaged in a hands-on comparison of models, can help us to understand what those details are.

ACKNOWLEDGEMENTS

This work was made possible by an award from Google.

REFERENCES

Abrams, D., Baecker, R., & Chignell, M. (1998). Information archiving with bookmarks: Personal web space construction and organization. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'98)*, 41-48, Los Angeles, CA.

Ames, M., & Naaman, M. (2007). Why we tag: Motivations for annotation in mobile and online media. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'07)*, 971-980, San Jose, CA.

Bälter, O. (1997). Strategies for organizing email messages. *Proceedings of HCI on People and Computers XII*, 21–38, Springer-Verlag, London.

Barreau, D., & Nardi, B. (1995). Finding and reminding: File organization from the desktop. *SIGCHI Bulletin, 27(3),* 39-43.

Bergman, O., Beyth-Marom R., & Nachmias, R. (2006). The project fragmentation problem in personal information management. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'06)*, 271-274, Montréal, Canada.

Bergman, O., et al. (2008). The user-subjective approach: A new direction for PIM systems design. *Third International CHI Workshop on Personal Information Management (PIM'08),* Florence, Italy. Retrieved 4/24/08 from http://pim2008.ethz.ch/papers/pim2008-bergman-etal.pdf

Boardman, R., & Sasse, M.A. (2004). Stuff goes in the computer but it doesn't come out: A cross-tool study of personal information management. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'04),* 583-590, Vienna, Austria.

Bondarenko, O., & Janssen, R. (2005). Documents at hand: Learning from paper to improve digital technologies. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'05)*, 121-130, Portland, OR.

Dourish, P., et al. (1999). Presto: An Experimental Architecture for Fluid Interactive Document Spaces. *ACM Transactions on Computer-Human Interaction, 6(2),* 133-161.

Golder, S., & Huberman, B.A. (2006).Usage patterns of collaborative tagging systems. *Journal of Information Science*, *32(2)*, 198-208.

Hearst, M.A. (2006). Clustering versus faceted categories for information exploration. *Communications of the ACM, 49(4)*, 59-61.

Jones, W. (2007). *Keeping Found Things Found: The Study and Practice of Personal Information Management*. San Francisco, CA: Morgan Kaufmann Publishers.

Jones, W., et al. (2005). Don't take my folders away!: Organizing personal information to get things done. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'05)*, 1505-1508, Portland, OR.

Jones, W., & Teevan, J. (2007). *Personal Information Management*. Seattle, WA: University of Washington Press.

Kipp, M.E.I., & Campbell, D.G. (2006). Patterns and inconsistencies in collaborative tagging systems: An examination of tagging practices. *Proceedings of the American Society for Information Science and Technology, 43(1)*, 178.

Kwasnik, B.H. (1991). The importance of factors that are not document attributes in the organization of personal documents. *Journal of Documentation, 47(4)*, 389-398.

Malone, T.W. (1983). How do people organize their desks?: Implications for the design of office information systems. *ACM Transactions on Information Systems, 1(1)*, 99-112.

Pak, R., Pautz, S., & Iden, R. (2007). Information organization and retrieval: An assessment of taxonomical and tagging systems. *Cognitive Technology, 12(1)*, 31-44.

Whittaker, S., & Sidner, C. (1996). Email overload: exploring personal information management of email. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'96)*, 276-283, Vancouver, Canada.